



Year: 2020

Does symptom severity matter in stepped and collaborative care for depression?

Watzke, Birgit ; Heddaeus, Daniela ; Steinmann, Maya ; Daubmann, Anne ; Wegscheider, Karl ; Härter, Martin

Abstract: **BACKGROUND:** We investigated the differential effectiveness of a stepped and collaborative care model (SCM) vs. treatment as usual (TAU) for primary care patients with various depression severity degrees and explored whether subgroups received distinct evidence-based treatments. **METHODS:** Sub-group analyses of a RCT were calculated applying a multiple linear mixed model with the factors 1. group (SCM; TAU), 2. severity ((mild-moderate (MMD); severe depression (SD)) and their interaction, with PHQ-9 as primary outcome. Utilization of treatments was analyzed descriptively. **RESULTS:** For the 737 participating patients (SCM: n = 569; TAU: n = 168), availability of data substantially varies between subgroups at 12-month follow-up ranging between 37% and 70%. ITT-analysis (Last-observation-carried-forward) revealed a significant interaction for group x severity [$p = 0.036$] and a significant difference between groups in symptom reduction for MMD (-3.9; [95% CI: -5.1 to -2.6, $p < 0.001$; $d = 0.64$] but not for SD (-1.6; [95% CI: -3.4 to 0.2, $p = 0.093$; $d = 0.27$]. Sensitivity analyses (multiple imputation, completer analysis, pattern mixture model) didn't confirm the interaction effect and showed significant effects for both severity groups with slightly higher effect sizes for MMD. Differences between SCM and TAU in the percentage of patients utilizing depression-specific treatments are larger for MMD. **LIMITATIONS:** There was a high proportion of missing values among severely depressed patients, especially in SCM. **CONCLUSION:** SCM is effective for both MMD and SD. Utilization patterns might help explain the higher effects for MMD. Various strategies of replacement of missing values lead to slightly divergent results due to selective drop out between severity groups.

DOI: <https://doi.org/10.1016/j.jad.2020.07.079>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-189854>

Journal Article

Accepted Version



The following work is licensed under a Creative Commons: Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) License.

Originally published at:

Watzke, Birgit; Heddaeus, Daniela; Steinmann, Maya; Daubmann, Anne; Wegscheider, Karl; Härter, Martin (2020). Does symptom severity matter in stepped and collaborative care for depression? *Journal of Affective Disorders*, 277:287-295.

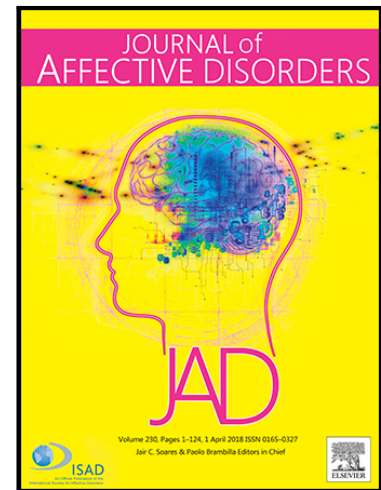
DOI: <https://doi.org/10.1016/j.jad.2020.07.079>

Journal Pre-proof

Does symptom severity matter in stepped and collaborative care for depression?

Prof. Dr. Birgit Watzke , Dr. Daniela Heddaeus , Maya Steinmann ,
Anne Daubmann , Prof. Dr. Karl Wegscheider ,
Prof. Dr. Martin Härter

PII: S0165-0327(20)32524-6
DOI: <https://doi.org/10.1016/j.jad.2020.07.079>
Reference: JAD 12249



To appear in: *Journal of Affective Disorders*

Received date: 25 January 2020
Revised date: 22 May 2020
Accepted date: 5 July 2020

Please cite this article as: Prof. Dr. Birgit Watzke , Dr. Daniela Heddaeus , Maya Steinmann ,
Anne Daubmann , Prof. Dr. Karl Wegscheider , Prof. Dr. Martin Härter , Does symptom severity matter
in stepped and collaborative care for depression?, *Journal of Affective Disorders* (2020), doi:
<https://doi.org/10.1016/j.jad.2020.07.079>

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2020 Published by Elsevier B.V.

Does symptom severity matter in stepped and collaborative care for depression?

Prof. Dr. Birgit Watzke^{1,a,*} b.watzke@psychologie.uzh.ch, Dr. Daniela Heddaeus^{2,a} d.heddaeus@uke.de, Ma-
ya Steinmann² m.steinmann@uke.de, Anne Daubmann³ a.daubmann@uke.de, Prof. Dr. Karl Wegscheider³
k.wegscheider@uke.uni-hamburg.de, Prof. Dr. Martin Härter² m.harter@uke.de

¹Clinical Psychology and Psychotherapy Research, Institute of Psychology, University of Zurich,
Binzmühlestrasse 14/16, CH-8050 Zurich/Switzerland

²Department of Medical Psychology, University Medical Center Hamburg-Eppendorf, Martinistraße 52,
20246 Hamburg, Germany

³Department of Medical Biometry and Epidemiology, University Medical Center Hamburg-Eppendorf, Mar-
tinistraße 52, 20246 Hamburg, Germany

*Corresponding author: Prof. Dr. Birgit Watzke Clinical Psychology and Psychotherapy Research, Institute of Psycholo-
gy, University of Zurich, Binzmühlestrasse 14/16, CH-8050 Zurich/Switzerland, Tel.: +41 44 635 73 21, Fax.: +41 44 635
73 29

^aequally contributed

Highlights

- SCM is effective for both mildly to moderately and severely depressed patients
- Effects are higher for mildly to moderately depressed patients
- Special attention must be paid to severely depressed patients
- SCM patients more often received outpatient depression-specific treatment
- Difference in treatment utilization is higher for not severely depressed patients

Abstract

Background: We investigated the differential effectiveness of a stepped and collaborative care model (SCM) vs. treatment as usual (TAU) for primary care patients with various depression severity degrees and explored whether subgroups received distinct evidence-based treatments.

Methods: Subgroup analyses of a RCT were calculated applying a multiple linear mixed model with the factor 1. group (SCM; TAU), 2. severity ((mild-moderate (MMD); severe depression (SD)) and their interaction, with PHQ-9 as primary outcome. Utilization of treatments was analyzed descriptively.

Results: For the 737 participating patients (SCM: n=569; TAU: n=168), availability of data substantially varies between subgroups at 12-month follow-up ranging between 37% and 70%. ITT-analysis (Last-observation-carried-forward) revealed a significant interaction for group x severity [$p=0.036$] and a significant difference between groups in symptom reduction for MMD (-3.9; [95% CI: -5.1 to -2.6, $p < 0.001$; $d=0.64$] but not for SD (-1.6; [95% CI: -3.4 to 0.2, $p = 0.093$; $d=0.27$]. Sensitivity analyses (multiple imputation, completer analysis, pattern mixture model) didn't confirm the interac-

tion effect and showed significant effects for both severity groups with slightly higher effect sizes for MMD. Differences between SCM and TAU in the percentage of patients utilizing specific treatments are larger for MMD.

Limitations: There was a high proportion of missing values among severely depressed patients, especially in SCM.

Conclusion: SCM is effective for both MMD and SD. Utilization patterns might help explain the higher effects for MMD. Various strategies of replacement of missing values lead to slightly divergent results due to selective drop out between severity groups.

Keywords

stepped care; collaborative care; outcome; depression; severity subgroups; treatment utilization

Abbreviations

SCM, Stepped and Collaborative Care Modell; TAU, Treatment as usual; RCT, Randomized controlled trial; MMD, mild to moderate depression; SD, severe depression; ITT, Intention to treat; CI, Confidence interval; ICD-10, International Statistical Classification of Diseases and Related Health Problems; DSM, Diagnostic and Statistical Manual of Mental Disorders; CBT, cognitive behavioral therapy; CCBT, computerized cognitive behavioral therapy; GP, General practitioner; LOCF, last-observation carried forward; MI, multiple imputation; AD, antidepressant treatment; PT, psychotherapeutic treatment

1. Introduction

Depression is one of the most common mental disorders. It is a very broad diagnostic category which can be divided into several subtypes of depression (DGPPN et al., 2015; Kessing, 2007; Kessler et al., 2003; NICE, 2010; Wittchen et al., 2000). The International Statistical Classification of Diseases and Related Health Problems (ICD-10) and the Diagnostic and Statistical Manual of Mental Disorders (DSM) differentiate mild, moderate and severe depression (American Psychiatric Association, 2013; Dilling et al., 2014). The ICD-10 additionally distinguishes between recurrent disorders and single episodes (Dilling et al., 2014). The differentiation into the subtypes of depression with regard to severity, duration and course is not only important for how depression presents clinically, but also for guideline-adherent decision-making regarding evidence-based treatments (DGPPN et al., 2015; NICE, 2010): international guidelines recommend providing active monitoring to patients with subthreshold depression or patients with mild depression as first step. Patients with mild to moderate depression or persistent subthreshold depressive symptoms should be offered low intensity treatments like bibliotherapy, individual guided self-help based on the principles of cognitive behavioral therapy (CBT), computerized cognitive behavioral therapy (CCBT), or physical activity programs. Either medication or psychotherapy should be offered to patients with mild to moderate depression or persistent subthreshold depressive symptoms who did not benefit from low intensity treatments. For severely depressed patients, a combination of medication and psychotherapy is recommended; in cases of severe and complex depression, this combined treatment should be offered regularly, sometimes also in an inpatient treatment setting if patients do not improve in outpatient care (DGPPN et al., 2015; NICE, 2010).

As there are distinct recommendations for different groups of depressed patients which vary in treatment intensity, a systematic treatment selection applying a stepped and collaborative care approach is crucial. While stepped care focuses on treating patients with an adequate treatment of the lowest possible intensity while continuously monitoring progress, collaborative care aims to systematically integrate different care providers. Both health care models are often combined and recommended by international depression guidelines (DGPPN et al., 2015; Landelijke Stuurgroep, 2005; NICE, 2010).

However, guideline implementation is still insufficient. Often patients with mild depression are over-prescribed with antidepressant medication (Baer et al., 2013; Trautmann and Beesdo-Baum, 2017), while patients with severe depression often do not receive the intensity of care they would need (Melchior et al., 2014). For the patients' sake but also from a cost-effectiveness perspective it would be relevant to prevent inpatient treatment by offering adequate outpatient treatment options at an early stage of treatment. These health care issues are addressed by guidelines and stepped and collaborative care models (SCM).

There is good evidence for the general effectiveness and partially also for the cost-effectiveness of SCM for depression ([Archer et al., 2012](#); [Firth et al., 2015](#); [Goorden et al., 2014](#); [Härter et al., 2018](#); [Ho et al., 2016](#); [van Straten et al., 2015](#); [Von Korff et al., 1997](#)). However, there is still little knowledge and evidence about differential effects of these models, i.e. whether there are characteristics of the patient or the type of the depression predicting the response to SCM. Since SCM comprises a complex intervention that is tailored individually for patients, in the differential analysis of effects treatment variables have also to be taken into account, i.e. the type of treatments actually received by patients. These analyses of individual treatment utilization and treatment pathways may help explain (variations in) outcome effects.

Even though it is known that depression severity degree is one of the best predictors of treatment response ([Katon et al., 1994](#); [Walker et al., 2000](#)), only few studies report results of SCM referring to different severity subgroups. One SCM intervention showed a significant improvement in depressive symptoms after 6 and 28 months compared with usual care for the group of moderately depressed patients, but not for patients in the high severity group ([Katon et al., 2002](#); [Walker et al., 2000](#)). In a SCM for depression in the Netherlands, Franx and colleagues found that patients in the non-severe depression group were treated according to the guideline oriented SCM protocol considerably more often than severely depressed patients. Furthermore, the adherence to the SCM improved during the study for the non-severe group, but not for the severe group. They concluded that special attention should be paid to the quality of care for severely depressed patients ([Franx et al., 2009](#)). However, another study reported that health care consumption increased significantly in both depressed patient groups ([Gidding et al., 2014](#)). In accordance with this, Unützer and colleagues did not find a significant interaction between intervention effect and depressive symptom severity ([Unützer et al., 2002](#)). In summary, no conclusive answer to this research question exists yet.

1.1. Aims of the study

Against this background, this article examines whether patients with different degrees of depression severity benefit differentially from stepped and collaborative care and whether they differ in the treatments received, taking into account severity degree. The general effectiveness of the model under investigation has been already reported elsewhere ([Härter et al., 2018](#)), while the present analyses combine subgroup analyses with the treatment utilization of the subgroups.

2. Methods

2.1 Study design

This parallel cluster-randomized controlled intervention trial assessed a consecutive sample of patients with depression from primary care with measurements at four time points (baseline and 3, 6 and 12 months after baseline). Randomization was conducted at the level of participating GP practices, dividing these into an intervention group and a control group in a 3:1 ratio. Patients recruited by GPs in the intervention group were treated in the stepped and collaborative care model (SCM), while patients recruited by GPs in the control group received treatment as usual (TAU). The process of randomization was conducted by a computer program using minimization based on practice size, practice location and income level of the practice's local district ([Watzke et al., 2014](#)).

Sample size calculation was based on the detection of a small to moderate effect (Cohen's d of 0.40) with a statistical power of 0.80 and a type I error rate of 0.05 (two sided) between SCM and TAU. After considering the clustered design with an assumed intra-cluster correlation of 0.05, the differential expected attrition rates between groups, and the plan to run further separate analyses for each of six treatment options, which are available in SCM (see section "2.5 Intervention"), we aimed to recruit a total of 860 patients. Expecting that every participating GP practice would recruit 15-25 patients, we planned to gain 40 practices for study participation (660 patients from 30 practices in SCM and 200 patients from 10 practices in TAU). In total, we recruited 49 practices: 36 practices were randomized to SCM and 13 practices to TAU ([Watzke et al., 2014](#)).

The study was approved by the Ethics Committee of the Hamburg Chamber of Psychotherapists and conducted according to the principles of the Declaration of Helsinki (2013 version). Written informed consent was obtained from all participants. The study was registered in ClinicalTrials.gov (registration number: NCT01731717; registration date: 11.12.2012).

2.2 Setting

The study “Health Network Depression” was embedded into the intersectoral research initiative *psychnet - The Hamburg Network for Mental Health* (2011-2015; funded by the German Federal Ministry of Education and Health) which consisted of 11 research and development projects on mental health care research in the metropolitan area of Hamburg (Germany) ([Härter et al., 2012](#); [Watzke et al., 2014](#)). Patient assessment and treatment in SCM took place within a multi-professional, specially established and trained network of GP practices, psychotherapists, psychiatrists and inpatient clinic representatives implemented in routine care ([Härter et al., 2015](#); [Heddaeus et al., 2015](#)), while assessment and treatment in TAU was conducted in GP practices as well as any facilities available in routine care.

2.3 Care Providers

All GP practices, psychotherapists, psychiatrists of the greater Hamburg area were invited by mail via the Hamburg Chamber of Physicians to participate in the study. Inclusion criteria for participation were willingness to participate in study procedures and to be working as a Statutory Health Insurance GP, psychotherapist or psychiatrist in an established practice. For inpatient institutions inclusion criteria were willingness to participate in study procedures and to have a license from the Association of Statutory Health Insurance.

2.4 Patients

Patient recruitment in SCM and TAU was carried out by the participating GPs and comprised three screening and assessment steps using checklists and the PHQ-9, successively ([Watzke et al., 2014](#)). Inclusion criteria were a minimum age of 18, five or more points on PHQ-9 and informed consent. Patients with insufficient knowledge of the German language or a health situation that did not allow questionnaire completion were excluded. Neither somatic nor mental comorbidities were exclusion criteria. Patients were only excluded if a mental disorder other than depression, e.g. a trauma-related disorder, was the main treatment focus. Patient inclusion took place from August 08, 2012 to March 31, 2014. Follow-up measurements took place between 2012 and 2015.

2.5 Interventions

2.5.1 Stepped collaborative care model (SCM)

The SCM is a stratified stepped and collaborative care approach which has been described in detail elsewhere ([Watzke et al., 2014](#)). GPs of the intervention group, i.e. GPs offering their patients SCM, completed an ICD-10-based checklist and imparted psycho-education if patients fulfilled the study criteria (PHQ-9 ≥ 5). Treatment interventions according to SCM were allocated by the GPs on the basis of the ICD-10 diagnosis following the evidence-based recommendations of the German National Clinical Practice Guideline “Unipolar Depression” ([Watzke et al., 2014](#)), i.e. criteria were: 1) depression severity and 2) patient preferences as result of a shared decision-making process.

Within the SCM, treatment options were available on four levels of intensity: **Step 1**) active monitoring; 2a) bibliotherapy ([Görlitz, 2010](#)); 2b) internet-based self-management ([Meyer et al., 2009](#)); 2c) telephone-administered psychotherapy (9-13 sessions) ([Steinmann et al., 2016](#)); 3a) outpatient cognitive-behavioral or psychodynamic psychotherapy (usually ≥ 25 sessions) as stand-alone treatment; 3b) antidepressant pharmacotherapy as stand-alone treatment; and 4) combination of psycho- and phar-

macotherapy in either out- or inpatient setting. Therefore, there were six treatment options ([step 2-4](#)) within the SCM – considering [Step 1](#) (“active monitoring”) as an antecedent option and not as actual treatment option.

Course of treatment and symptoms was monitored by care providers with the PHQ-9 in predefined intervals depending upon intervention and symptoms. If the PHQ-9 score had not improved by at least 20% since beginning a treatment, stepping up was recommended. A care provider network was built up to facilitate communication and prompt referral between primary and secondary care. To improve referral and to reduce waiting times, an online platform indicating available treatment capacities was implemented. All participating multi-professional providers obtained intensive training regarding the relevant guideline recommendations (DGPPN et al., [2015](#); [NICE, 2004](#)), the SCM and its related interventions. Quarter-yearly quality circles took place to assure the quality of the integrated care model and a stable adherence to clinical recommendations.

2.5.2 Treatment as usual (TAU)

Patients in the TAU received treatment as usual by their GP and within the regular German health care system, including potentially necessary referrals to psychotherapy and psychiatry in out- or inpatient care settings. Systematic screening was carried out in both SCM and TAU to ensure a comparable recruitment and inclusion process.

2.6 Outcomes

All outcomes were assessed by self-report questionnaires at four time points: The baseline (T0) assessment was handed out by the GP at study inclusion and filled out before any treatment began. The questionnaires 3 (T1), 6 (T2) and 12 months (T3) after baseline were sent to patients by mail. If a questionnaire was not returned within two weeks, up to two reminder letters were sent to the patient and one to the patient's GP within the following month to improve response rates.

Primary outcome was change in depressive symptoms assessed by the Patient Health Questionnaire-9 (PHQ-9) ([Löwe et al., 2004](#)) from baseline to 12 months (differential treatment outcome). Treatment utilization during the 12 months of study participation was assessed by patient self-report questionnaires at T1, T2 and T3. Patients retrospectively reported their use of different treatment options and consultation of various care providers.

2.7. Statistical analysis

The outcome analysis regarding change in PHQ-9 from baseline to T3 was based on the intention to treat (ITT) population. Only this time point was part of our analysis. In case of missing follow-up values, a last-observation carried forward (LOCF) imputation was performed using also the time points in between for the imputation (primary analysis) ([Watzke et al., 2014](#)). As drop-outs were distributed unequally, three kinds of sensitivity analyses were performed: One model using the ITT population and the multiple imputation (MI) replacement method instead of LOCF and a second one without LOCF replacement but with only those cases who provided data for the main outcome at T3 (i.e. completer analyses). To account for a possible violation of the missing at random assumption, we used a pattern mixture model with control-based pattern imputation as a third sensitivity analysis ([Ratitch et al., 2013](#)).

For the subgroup analysis, we divided patients in both conditions into two groups of different depression severity: mild to moderate depression severity (averaged PHQ-9-score of less than 15) and severe depression severity (averaged PHQ-9-score of 15 or more). We calculated a multiple linear mixed model with group (SCM vs TAU), depression severity (mild-moderate vs. severe), their interaction and education level as fixed effects (due to differences at baseline in this variable, cp. [Table 1](#)) and the general practice care unit as a random effect. We fitted a random intercept model with a scaled identity covariance matrix. We reported adjusted marginal means of the interaction term and their 95% confidence interval and p values. The significance level was set at 5% (two-sided).

The differential treatment utilization of the two groups was analyzed descriptively. The analysis of treatment utilization is based on the sample of patients for which T3 data (including information on treatment utilization during the whole observation period of 12 months) is available (see also 3.3.2). The following hypotheses regarding differences in treatment utilization were derived from the stepped care model: In both severity groups (MMD und SD) it was expected that SCM was associated with fewer patients being treated with nothing besides GP consultations, as well as to a higher utilization of depression-specific treatments. For MMD, SCM was expected to result in the utilization of low intensity treatments, more psychotherapeutic treatment, less anti-depressant (AD) treatment (as AD is not the first line treatment for mild depression) and less inpatient treatment than TAU. For severe depression, SCM was expected to be associated with more psychotherapeutic treatment (PT), more AD treatment, and more combination of PT and AD.

All analyses were conducted with SPSS 23 and SAS 9.4.

3. Results

3.1. Participant flow

[Figure 1](#) shows the participant flow. 737 patients were included (SCM: n = 569 vs. TAU: n = 168), with data available for 60% of SCM patients and 64% of TAU patients after 12 months. In SCM, 67% of patients (n=383) had a PHQ-9 score of less than 15 points (mild to moderate depression (MMD)). In TAU this was the case for 67% (n=113). The severe depression group (SD) in SCM consisted of n=186 (33%) and in TAU of n=55 patients (33%). After baseline assessment, 46% of SD patients in SCM and 31% in TAU dropped out. For MMD, this rate was lower, with 11% of patients in SCM and 14% in TAU dropping out. At T3, 70% of the data of the MMD group were available in both SCM and TAU. In the SD group, 53% of the data were available in TAU and 37% in SCM.

3.2. Sample

[Table 1](#) shows the baseline characteristics of patients in the four subgroups.

3.3. Differential treatment outcome

3.3.1. Primary analysis

The LOCF-analysis of the adjusted PHQ-9 mean reduction from baseline to 12 months revealed a significant interaction between group and depression severity [$p=0.036$] ([Table 2](#)). Symptom reduction on the PHQ-9 for MMD was -6.2 points in the SCM and -2.4 points in TAU. For SD, symptom reduction on the PHQ-9 was -2.6 in SCM and -1.0 points in TAU. With a difference of 3.9 points in the PHQ-9, the MMD group in SCM showed a significantly greater reduction in PHQ-9 scores than patients in TAU [95% confidence interval (CI): -5.1 to -2.6, $p < 0.001$; Cohen's $d=0.62$]. For the SD group, the difference between SCM and TAU (a symptom reduction of 1.6 points) was not significant [95% CI: -3.4 to 0.3, $p = 0.093$; Cohen's $d=0.25$].

3.3.2 Sensitivity analyses

[Table 2](#) also summarizes the results of the three different model calculations for the sensitivity analyses (ITT analyses with MI replacement method and completer analyses and pattern mixture model with control-based pattern imputations). For all forms of analyses, an interaction between group and severity was not confirmed. The difference in symptom reduction between SCM and TAU was significant in both severity groups, but with higher effect sizes for MMD than for SD.

3.4. Differential treatment utilization

[Table 3](#) summarizes the observed proportions of treatments patients received during the 12 months of study participation and the comparison between SCM and TAU in total and differentiated for severity groups.

GP consultation. As recruitment took place in primary care, nearly all patients in SCM and TAU were seen by their GP (90.6 % and 96.3 % respectively). However, patients with TAU more often received nothing besides GP consultation, compared to patients in SCM (20.4 % vs. 3.5 %). Differentiated by severity subgroups, it becomes clear that this difference can be traced to the MMD group (25.3 % TAU patients vs. 3.7% SCM patients with GP consultation only), whereas for the SD group the percentage is low in both conditions.

Any outpatient depression-specific treatment. The utilization of outpatient depression-specific interventions (a global category summarizing low intensity-treatments, psychotherapy and pharmacotherapy) is in line with this. Here, there is a different utilization pattern of 16.9% between TAU and SCM. Differentiated by severity groups, this difference can also be traced to the MMD group (72.2% TAU patients vs. 94.4% SCM patients with depression-specific treatment), whereas for the SD group more than 93% of the patients in both conditions received this kind of treatment. This pattern results in a 20% greater difference between SCM and TAU for MMD than for SD (22.2% vs. 2.6%).

Low intensity treatments. Low intensity treatments which were only offered in SCM (telephone psychotherapy, internet self-help, and bibliotherapy) were utilized by more than 40% of SCM patients (43.7% in MMD and 30.4% in SD). 20.7% of MMD patients utilized this form of intervention exclusively, i.e. without receiving a more intensive treatment in addition. In the SD group, there is a percentage of 15.9% of patients utilizing only low intensity treatments.

Psychotherapy. For the global category, i.e. psychotherapy alone or in combination, there is a difference in utilization of 13.4% between TAU and SCM. Differentiating for severity groups, this difference is higher for MMD (16.6%) than for SD (6.0%): Accordingly, psychotherapy is initiated for 45.6% of MDD patients in TAU, while in the other three groups the utilization rate is 62% or higher. However, confidence intervals for these comparisons (total and MMD patients) overlap slightly.

For the subgroup of patients receiving psychotherapy only and for the subgroup receiving psychotherapy and antidepressant medication, no differences in utilization between groups was found. Since there is no difference for these treatment utilization patterns, it can be concluded that the difference for the global (category psychotherapy alone or in combination) is due to the use of psychotherapy in combination with low-intensity treatments.

Treatment with antidepressant pharmacotherapy. AD as global category, i.e. alone or in combination, differences in utilization between groups cannot be confirmed. Approximately half of the patients with MMD received AD (54.4% in TAU and 49.6% in SCM), whereas more than 2/3 of patients with SD received AD (82.8% in TAU and 73.9% in SCM).

For the subgroups of patients receiving AD only there is a difference in utilization between TAU and SCM: This difference is similar across severity groups (MMD: 20.3% TAU vs. 5.9% SCM; SD: 17.2% TAU vs. 7.2% SCM patients with AD only). However, due to the large and overlapping confidence interval in the SD group the difference is reliable only for the MMD group.

Inpatient treatment. Inpatient care was initiated 12.6% less frequently in SCM than in TAU (21.5% TAU vs. 8.9% SCM patients with inpatient treatment). For the SD group, approx. 1/3 of patients in both conditions received inpatient care (39.1% TAU and 34.5% SCM patients).

5. Discussion

5.1 Differential treatment outcome

This subgroup analysis of a cluster-randomized controlled intervention trial suggests that SCM is effective for patients of different severity groups. Our primary strategy for data analysis (ITT analysis with replacement of missing values by the LOCF method) shows a significant effect regarding depressive symptom reduction for SCM in mildly to moderately depressed patients but not in severely depressed patients. However, the sensitivity analyses (ITT-analysis with MI replacement, the completer analysis and pattern mixture model with control-based pattern imputation) do not confirm this result: According to these three sensitivity analyses, there are significant effects for SCM in both severity groups. However, effect sizes for SCM are higher for mildly to moderately depressed patients compared to severely depressed patients. Given these slightly divergent results, a relevant proportion of the differential effect of SCM for different severity groups in the LOCF analysis might be due to the high rate of selective dropouts in the group of severe depression patients, i.e. especially those treated in the SCM. A closer look at the dropouts in severely depressed patients reveals that most of them left the study immediately after study inclusion and did not contribute any other data except baseline assessment - neither regarding patient questionnaires nor the documentation data of their GPs. Replacing missing values by LOCF, i.e. by the baseline score, for a substantial percentage (46%) within the severely depressed patients in SCM might have led to an underestimation of effectiveness for this subgroup. Therefore, defining LOCF as primary strategy of replacement of missing values when designing the study turns out to be unfavourable in the course of the trial revealing the (unexpected) the selective drop out between severity groups. Accordingly, when changing the replacement strategy and choosing MI replacement or pattern mixture model, SCM is also effective for severe depression. The differential effect sizes of stepped and collaborative care models with larger effects for mild to moderate patients are in line with those few outcome studies taking into account symptom severity of depression ([Katon et al., 2002](#); [Walker et al., 2000](#)). Interestingly, this differential effect points in the reverse direction compared to results of meta analyses on placebo-controlled antidepressant trials suggesting an effect of AD only with increasing severity of depression symptoms ([Fournier et al., 2010](#); [Khan et al., 2002](#); [Rabinowitz et al., 2016](#)). However, this is not contradictory to the results of our study as we have not investigated a single treatment options such as antidepressant medication but a whole care model aiming at tailoring treatments to patients – and relying the treatment selection on guideline recommendations which are exactly based on the evidence on differential effects of antidepressant medication (among others). This has led to the algorithm within the SCM that in mild to moderate depression the first line treatment was (low intensity) psychological treatment - and not antidepressant medication.

What might have caused the selective dropout in the severely depressed group in our study? We do not have any information about the reasons for this. The severity of these patients' depression and the associated symptom burden may have impeded their participation in data collection and staying in continuous contact with the responsible care providers, as SCM recommends. Franx et al. report similar difficulties in adherence for patients with severe depression participating in stepped care ([Franx et al., 2009](#)). When severely depressed patients joined the SCM and were treated according to its concept, they benefitted significantly more from treatment in SCM than patients in TAU, as shown in the completer analysis.

However, even when taking into account the methodical uncertainty caused by the selective dropout, we did find - at least slightly - lower effect sizes for severely depressed patients in both SCM and TAU implicating that this group profits less from treatment than mildly to moderately depressed patients. This could be confirmed by all analyses (again with variation across primary analysis on the one hand and sensitivity analyses on the other hand) and is in line with findings of other studies that report better outcomes for non-severely depressed patients than for severely depressed patients ([Franx et al., 2009](#); [Katon et al., 2002](#); [Walker et al., 2000](#)).

5.2 Differential treatment utilization

The analysis of treatments received by patients during 12-month study participation provides important information for understanding the results of the differential outcomes and allows an insight into which treatments and interventions were actually implemented in SCM and routine care in the different subgroups. Overall, the results of the treatment utilization analysis are consistent with those of the effectiveness analysis: they help explain the higher benefit of the SCM for mildly to moderately depressed patients in terms of adequate treatment allocation. The hypothesis that patients in SCM more often receive depression-specific treatment and less often only GP consultations was confirmed. Additionally, a higher difference in these treatment utilization patterns and therefore a higher impact of the SCM on utilization was observed for mildly to moderately depressed patients in

comparison to severely depressed patients. This finding can partly be explained by the successful implementation of low intensity approaches as stand-alone treatment for mildly to moderately depressed patients. These treatments were additional intervention options GPs in SCM could use. However, this stand-alone application of low intensity treatments is adherent to the SCM concept only for mildly to moderately depressed patients, but not for severely depressed patients. However, the latter were also provided with only this treatment in 16% of the cases. Thus, these severely depressed patients were not treated in line with the SCM rational and with guidelines, and it can be assumed that they were undersupplied: for this group, the low intensity treatment did not facilitate the access to more intense treatments by introducing them into the health care system and motivating them for further treatment (vs. approx. 14% severely depressed patients who received an additional specific treatment, i.e. for whom the low intensity treatment succeeded in serving as “door opener”). This may explain the lower effect size of the effectiveness analysis in the SCM for the subgroup of severely depressed patients, at least to a certain degree. The hypothesis of reduced AD use in SCM for mildly to moderately depressed patients in general could not be confirmed. The reason for this might be that the group of mild to moderate depression does not only consist of mildly but also of moderately depressed patients, for whom it would be adequate to apply AD treatment as alternative treatment option to PT. However, there were less mildly to moderately depressed patients in SCM who received nothing but AD treatment, i.e. these patients received additional psychological support or used a self-help approach. The hypothesis of a reduced proportion of inpatient treatments for mildly to moderately depressed patients in SCM was confirmed, as admission rates in SCM were less than half of those in TAU (9% vs. approx. 22%). This shows that the SCM offers the potential to avoid this intensive and also cost-intensive care form for patients with non-severe depression, while at the same time improving the outcome of this group, as effectiveness analysis show.

The utilization data suggests that the focus of the GP in TAU lay on severely depressed patients while mildly to moderately depressed patients were more likely overlooked. This was different in SCM and implicates that the SCM holds potential to detect and improve the undersupply of mildly to moderately depressed patients.

Again, as we refer to the same assessment in the analysis of utilization as in the outcome analyses (assessment at T3), the data base for utilization is broader for the mildly to moderately depressed group than for the severely depressed group, given the larger dropout rate of the latter. Therefore, comparisons within the severe depression group are less reliable due to the relatively small sample size and associated larger confidence intervals. However, referring to the patients who participated the whole study period of 12 months (completers) we were able to robustly report patterns of utilization.

Limitations. The major limitation of these outcome and utilization analyses is the aforementioned high proportion of missing values in the subgroup of severely depressed patients. Therefore, our analyses are more robust for the mild to moderate than for the severely depressed group. In order to respond to this limitation we conducted sensitivity analyses in order to take into account this circumstance. A limitation of the analysis of utilization is that due to the kind of data assessment (i.e. retrospectively assessed utilization data reported by patients) it is not possible to deduce statements about the exact temporal sequence of treatments, stepping-up pattern and – especially important – the quality of the treatments patients received, e.g. the quality of type, dose and frequency of medication in pharmacotherapy.

5.3 General discussion

As both the differential outcome analyses with lower effect sizes for severely depressed patients in SCM and the treatment utilization analysis show, the particular challenge is to improve the care of severely depressed patients and to ensure more guideline-based care. Cases of severe depression are more complex and need more effort to implement guideline-adherent treatment. The results of our study suggest that certain interventions of the SCM especially for severe depression (e.g. combination of PT and AD) have not yet been implemented sufficiently and that there is need to improve the implementation of the SCM in line with its concept. A prior study analyzing the SCM patients of the same trial regarding treatment decisions and initial treatments reported consistent results: guideline-adherent treatment initiation was highest for mildly depressed patients and lowest for severely depressed patients, and the latter more often did not receive the treatment decided upon with their GP ([Heddaeus et al., 2018](#)).

The higher barriers in care delivery for severely depressed patients might be due to their higher symptom burden, which may affect treatment implementation. For example, the challenge to call a psychotherapist for an appointment may be higher for a severely depressed patient than for a mildly depressed patient who suffers from less symptoms of social withdrawal and hopelessness. Patients in severe crises might also be more ambivalent towards treatments and decide against a selected treatment even before beginning it, due to doubts and negative expectations and emotions. One implication of this finding for the improvement of depression care could be the comprehensive implementation of closer systematic monitoring by the main care provider to check whether the patient receives the agreed-upon treatment and whether this leads to relevant symptom reduction or must be adapted. Another option would be the introduction of an even more systematic case management, which tracks and accompanies the patient throughout his or her treatment pathway. The conception of stepped and collaborative care imply this kind of support – however, implementing this properly within the complex intervention remains a crucial task.

6. Conclusions

In conclusion, access to treatment, triage and the coordination of pathways remains an area with challenges for which stepped care promises to offer a means of improvement for different severity levels of depression. An adequate implementation of the approach demands much effort by providers, patients and stakeholders (the latter for creating the adequate conditions and incentives within the health care system) and appears to be crucial for ensuring processes and – by this – improved outcome for patients.

7. Author Contributions

M.H. and B.W. conceptualised and designed the study, obtained funding and were responsible for its conduct and overall supervision. A.D., K.W., H.H.K. and C.B. contributed to specific methodical and health economic issues. M.S., D.H., S.L., B.W. and M.H. organised the care provider network, recruitment process and data collection and management. M.H., B.W., A.D., K.W., M.S. and D.H. were actively involved in analysis and interpretation of data. All authors contributed, read and approved the final manuscript.

8. Additional information

Department of Medical Psychology, University Medical Centre Hamburg-Eppendorf, Germany

Contributors

BW, DH and MH conceptualised and designed the analyses and were responsible for its conduct. BW and MH were responsible for the overall supervision. AD and KW contributed to specific methodical issues. DH, BW, AD, MS and AD were actively involved in the analysis and interpretation of data. All authors contributed, read and approved the final manuscript.

Data availability statement

The datasets generated and analyzed during the current study are available from the study center on reasonable request.

Uncited References:

[GGZ, 2005](#)

Comment [AU1]: CE: Please check Contributors and Data availability statement

Comment [AU2]: This section comprises references that occur in the reference list but not in the body of the text. Please position each reference in the text or, alternatively, delete it.

Declaration of Competing Interest

None

Acknowledgements

The study was funded within the larger project psychenet – The Hamburg Network for Mental Health. psychenet was a project network funded by the German Federal Ministry of Education and Research (funding code 01KQ1002B) in the region of Hamburg which consists of more than 80 scientific and medical institutions, counseling centers, the Senate and the Chamber of Commerce of the Free and Hanseatic City of Hamburg, companies, as well as patients' and relatives' associations (2011 – 2015; www.psychenet.de). We would like to thank all care providers and patients who participated in this study.

Funding: This work was supported by the German Federal Ministry of Education and Research (grant number 01KQ1002B)

References

- < bib id="bib1" label="Archer et al., 2012" type="Periodical">Archer, J., Bower, P., Gilbody, S., Lovell, K., Richards, D., Gask, L., Dickens, C., & Coventry, P. (2012). Collaborative care for depression and anxiety problems. *Cochrane database of systematic reviews (Online)*, 10, CD006525.</bib>
- < bib id="bib2" label="American Psychiatric Association 2013" type="Book">American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders, DSM-V*. Washington: American Psychiatric Association.</bib>
- < bib id="bib3" label="Baer et al., 2013" type="Periodical">Baer, N., Schuler, D., Füglistner-Dousse, S., & Moreau-Gruet, F. (2013). *Depressionen in der Schweiz. Daten zur Epidemiologie, Behandlung und sozial-beruflichen Integration*. Retrieved from Neuchâtel:</bib>
- < bib id="bib4" label="DGPPN, BÄK, KBV, AWMF, AkdÄ, BpTK, BApK, DAGSHG, DEGAM, DGPM, DGPs, & DGRW (Eds.) 2015" type="Book">DGPPN, BÄK, KBV, AWMF, AkdÄ, BpTK, BApK, DAGSHG, DEGAM, DGPM, DGPs, & DGRW (Eds.). (2015). *S3-Leitlinie/Nationale VersorgungsLeitlinie Unipolare Depression (2. Auflage, Version 2 ed.)*. Berlin: DGPPN, BÄK, KBV, AWMF.</bib>
- < bib id="bib5" label="Dilling et al., 2014" type="Book">Dilling, H., Mombour, W., & Schmidt, M. H. (2014). *Internationale Klassifikation psychischer Störungen: ICD-10 Kapitel V (F) - Klinisch-diagnostische Leitlinien*. Bern: Verlag Hans Huber.</bib>
- < bib id="bib6" label="Firth et al., 2015" type="Periodical">Firth, N., Barkham, M., & Kellett, S. (2015). The clinical effectiveness of stepped care systems for depression in working age adults: A systematic review. *J Affect Disord*, 170, 119-130. doi:<http://dx.doi.org/10.1016/j.jad.2014.08.030></bib>
- < bib id="bib7" label="Fournier et al., 2010" type="Periodical">Fournier, J., DeRubeis, R., Hollon, S., Dimidjian, S., Amsterdam, J., Shelton, R., & Fawcett, J. (2010). Antidepressant drug effects and depression severity: a patient-level meta-analysis. *JAMA*, 303, 47-53.</bib>
- < bib id="bib8" label="Franx et al., 2009" type="Periodical">Franx, G., Meeuwissen, J. A. C., Sinnema, H., Spijker, J., Huyser, J., Wensing, M., & de Lange, J. (2009). Quality improvement in depression care in the Netherlands: the De-

pression Breakthrough Collaborative. A quality improvement report. *Int J Integr Care [Electronic Resource]*, **9**, e84. </bib>

<bib id="bib9" label="GGZ, 2005" type="Periodical">GGZ, L. S. M. R. i. d. (2005). *[Multidisciplinary guideline depression: guideline for diagnostics and treatment of adult clients with a major depressive disorder 2005]*. Utrecht. </bib>

<bib id="bib10" label="Gidding et al., 2014" type="Periodical">Gidding, L. G., Spigt, M. G., & Dinant, G.-J. (2014). Stepped collaborative depression care: primary care results before and after implementation of a stepped collaborative depression programme. *Fam Pract*, **31**(2), 180-192. </bib>

<bib id="bib11" label="Goorden et al., 2014" type="Periodical">Goorden, M., Muntingh, A., van Marwijk, H., Spinhoven, P., Ader, H., van Balkom, A., van der Feltz-Cornelis, C., & Hakkaart-van Roijen, L. (2014). Cost utility analysis of a collaborative stepped care intervention for panic and generalized anxiety disorders in primary care. *J Psychosom Res*, **77**(1), 57-63. </bib>

<bib id="bib12" label="Görlitz, 2010" type="Book">Görlitz, G. (2010). *[Self-help for Depression]*. Stuttgart: Klett-Cotta. </bib>

<bib id="bib13" label="Härter et al., 2015" type="Periodical">Härter, M., Heddaeus, D., Steinmann, M., Schreiber, R., Brettschneider, C., König, H.-H., & Watzke, B. (2015). [Collaborative and stepped care for depression. Development of a model project within the Hamburg Network for Mental Health (psychnet.de)]. *Bundesgesundheitsbl Gesundheitsforsch Gesundheitsschutz*, **58**(4-5), 420-429. doi:10.1007/s00103-015-2124-7 </bib>

<bib id="bib14" label="Härter et al., 2012" type="Periodical">Härter, M., Kentgens, M., Brandes, A., Bock, T., Dirmaier, J., Erzberger, M., Furstenberg, W., Hillebrandt, B., Karow, A., Von Dem Knesebeck, O., König, H.-H., Löwe, B., Meyer, H.-J., Romer, G., Rouhiainen, T., Scherer, M., Thomasius, R., Watzke, B., Wegscheider, K., & Lambert, M. (2012). Rationale and content of psychnet: The Hamburg Network for Mental Health. *Eur Arch Psy Clin N*, **262**(2 SUPPL.), 57-63. </bib>

<bib id="bib15" label="Härter et al., 2018" type="Periodical">Härter, M., Watzke, B., Daubmann, A., Wegscheider, K., HH, K., Brettschneider, C., Liebherz, S., Heddaeus, D., & Steinmann, M. (2018). Guideline-based stepped and collaborative care for patients with depression in a cluster-randomised trial. *Sci Rep-UK*, **8**, 9389. doi:10.1038/s41598-018-27470-6 </bib>

<bib id="bib16" label="Heddaeus et al., 2018" type="Periodical">Heddaeus, D., Steinmann, M., Daubmann, A., Härter, M., & Watzke, B. (2018). Treatment selection and treatment initialization in guideline-based stepped and collaborative care for depression. *PLoS ONE*, **13**(12), e0208882. doi:10.1371/journal.pone.0208882 </bib>

<bib id="bib17" label="Heddaeus et al., 2015" type="Periodical">Heddaeus, D., Steinmann, M., Liebherz, S., Härter, M., & Watzke, B. (2015). [psychnet - The Hamburg Network for Mental Health: Evaluation of the Health Network Depression from the Perspective of Participating General Practitioners, Psychotherapists and Psychiatrists]. *Psychiatr Prax*, **42**(Supplement 1), 51-56. </bib>

<bib id="bib18" label="Ho et al., 2016" type="Periodical">Ho, F. Y., Yeung, W. F., Ng, T. H., & Chan, C. S. (2016). The Efficacy and Cost-Effectiveness of Stepped Care Prevention and Treatment for Depressive and/or Anxiety Disorders: A Systematic Review and Meta-Analysis. *Sci Rep*, **6**, 29281. doi:10.1038/srep29281 </bib>

- < bib id="bib19" label="Katon et al., 1994" type="Periodical">Katon, W., Lin, E., von Korff, M., Bush, T., Walker, E., Simon, G., & Robinson, P. (1994). The predictors of persistence of depression in primary care. *J Affect Disord*, 31(2), 81-90. </bib>
- < bib id="bib20" label="Katon et al., 2002" type="Periodical">Katon, W., Russo, J., Von Korff, M., Lin, E., Simon, G., Bush, T., Ludman, E., & Walker, E. (2002). Long-term Effects of a Collaborative Care Intervention in Persistently Depressed Primary Care Patients. *J Gen Intern Med*, 17(10), 741--748. doi:10.1046/j.1525-1497.2002.11051.x</bib>
- < bib id="bib21" label="Kessing, 2007" type="Periodical">Kessing, L. (2007). Epidemiology of subtypes of depression. *Acta Psychiatr Scand*, 115(433), 85-89. </bib>
- < bib id="bib22" label="Kessler et al., 2003" type="Periodical">Kessler, R. C., Berglund, P., Demler, O., Jin, R., Koretz, D., Merikangas, K. R., Rush, A. J., Walters, E. E., & Wang, P. S. (2003). The epidemiology of major depressive disorder: Results from the National Comorbidity Survey Replication (NCS-R). *JAMA*, 289(23), 3095-3105. </bib>
- < bib id="bib23" label="Khan et al., 2002" type="Periodical">Khan, A., Leventhal, R., Khan, S., & Brown, W. (2002). Severity of depression and response to antidepressants and placebo: an analysis of the Food and Drug Administration database. *J Clin Psychopharmacol*, 22(1), 40-45. </bib>
- < bib id="bib24" label="Löwe et al., 2004" type="Periodical">Löwe, B., Unützer, J., Callahan, C. M., Perkins, A. J., & Kroenke, K. (2004). Monitoring depression treatment outcomes with the patient health questionnaire-9. *Med Care*, 42(12), 1194-1201. </bib>
- < bib id="bib25" label="Melchior et al., 2014" type="Periodical">Melchior, H., Schulz, H., & Härter, M. (2014). *Faktencheck Gesundheit: Regionale Unterschiede in der Diagnostik und Behandlung von Depressionen*. Retrieved from Gütersloh: </bib>
- < bib id="bib26" label="Meyer et al., 2009" type="Periodical">Meyer, B., Berger, T., Caspar, F., Beevers, C. G., Andersson, G., & Weiss, M. (2009). Effectiveness of a novel integrative online treatment for depression (Deprexis): Randomized controlled trial. *J Med Internet Res*, 11(2), e15. doi:10.2196/jmir.1151 </bib>
- < bib id="bib27" label="NICE 2004" type="Book">NICE. (2004). *Depression: Management of depression in primary and secondary care (Clinical guideline 22)*. London: National Institute for Clinical Excellence. </bib>
- < bib id="bib28" label="NICE 2010" type="Book">NICE. (2010). *Depression. The treatment and management of depression in adults (updated edition)*. Leicester: The British Psychological Society. </bib>
- < bib id="bib29" label="Rabinowitz et al., 2016" type="Periodical">Rabinowitz, J., Werbeloff, N., Mandel, F., Menard, F., Marangell, L., & Kapur, S. (2016). Initial depression severity and response to antidepressants v. placebo: patient-level data analysis from 34 randomised controlled trials. *Brit J Psychiat*, 209(5), 427-428. </bib>
- < bib id="bib30" label="Ratitch et al., 2013" type="Periodical">Ratitch, B., O'Kelly, M., & Tosiello, R. (2013). Missing data in clinical trials: From clinical assumptions to statistical analysis using pattern mixture models. *Pharmaceutical Statistics*, 12(6), 337-347. doi:https://doi.org/10.1002/pst.1549 </bib>
- < bib id="bib31" label="Steinmann et al., 2016" type="Periodical">Steinmann, M., Heddaeus, D., Härter, M., Lieberherz, S., Weymann, N., & Watzke, B. (2016). [Telephone-administered cognitive behavioral psychotherapy for depression: A German treatment program]. [Telephone-administered cognitive behavioral psychotherapy for depression: A German treatment program]. *Verhaltenstherapie*, 26, 244-253. doi:10.1159/000452435 </bib>

- <bib id="bib32" label="Trautmann and Beesdo-Baum, 2017" type="Periodical">Trautmann, S., & Beesdo-Baum, K. (2017). The Treatment of Depression in Primary Care. *Dtsch Arztebl Int*, 114(43), 721-728. doi:10.3238/arztebl.2017.0721 </bib>
- <bib id="bib33" label="Unützer et al., 2002" type="Periodical">Unützer, J., Katon, W., Callahan, C., Williams, J., Hunkeler, E., Harpole, L., Hoffing, M., Della Penna, R., Noel, P. H., Lin, E., Arean, P., Hegel, M. T., Tang, L., Belin, T. R., Oishi, S., & Langston, C. (2002). Collaborative Care Management of Late-Life Depression in the Primary Care Setting: A Randomized Controlled Trial. *JAMA*, 288(22), 2836-2845. </bib>
- <bib id="bib34" label="van Straten et al., 2015" type="Periodical">van Straten, A., Hill, J., Richards, D. A., & Cuijpers, P. (2015). Stepped care treatment delivery for depression: a systematic review and meta-analysis. *Psychol Med*, 45(2), 231-246. </bib>
- <bib id="bib35" label="Von Korff et al., 1997" type="Periodical">Von Korff, M., Katon, W., Bush, T., Lin, E. H. B., Simon, G. E., Saunders, K., Ludman, E., Walker, E., & Unutzer, J. (1997). Treatment costs, cost offset, and cost effectiveness of collaborative management of depression. *Psychosomatic Medicine*, 60, 143-149. </bib>
- <bib id="bib36" label="Walker et al., 2000" type="Periodical">Walker, E., Katon, W., Russo, J., MR, v. K., Lin, E., Simon, G., Bush, T., Ludman, E., & Unützer, J. (2000). Predictors of outcome in a primary care depression trial. *J Gen Intern Med*, 15, 859-867. </bib>
- <bib id="bib37" label="Watzke et al., 2014" type="Periodical">Watzke, B., Heddaeus, D., Steinmann, M., König, H.-H., Wegscheider, K., Schulz, H., & Härter, M. (2014). Effectiveness and cost-effectiveness of a guideline-based stepped care model for patients with depression: Study protocol of a cluster-randomized controlled trial in routine care. *BMC Psychiatry*, 14, 230. </bib>
- <bib id="bib38" label="Wittchen et al., 2000" type="Periodical">Wittchen, H.-U., Müller, N., Schmidt-kunz, B., Wittchen, S., & Pfister, H. (2000). Erscheinungsformen, Häufigkeit und Versorgung von Depressionen. *Fortschritte der Medizin Originalien*, 118. Jg., Sonderheft 1/2000. </bib>

Figure 1: Participant flowchart according to the cluster-randomized CONSORT statement (40). Legend: GP: General practitioner; MMD: mild to moderate depression; SD: severe depression; AFA: available for analysis; LOV: last observed value; ITT: Intention-to-treat

Table 1: Baseline characteristics on patient level in SCM₁ and TAU₂

Baseline characteristics	Group			
	Mild-moderate		Severe	
	SCM ₁ (n=383)	TAU ₂ (n=113)	SCM ₁ (n=186)	TAU ₂ (n=55)
Age (M (Standard Deviation))	43.4 (13.9)	48.9 (16.4)	39.4 (12.2)	38.9 (10.8)
Female gender (% (n))	75.2 (288)	77.9 (n=88)	66.7 (124)	72.7 (40)

Education level (% (n))

<i>Secondary general school</i> ₃	18.8 (72)	32.7 (37)	23.7 (44)	23.6 (13)
<i>Intermediate secondary school</i> ₄	29.0 (111)	23.0 (26)	24.7 (46)	32.7 (18)
<i>High school</i> ₅	25.8 (99)	23.9 (27)	23.7 (44)	18.2 (10)
<i>University or technical college degree</i>	17.5 (67)	10.6 (12)	7.5 (14)	9.1 (5)
<i>No school degree</i>	1.6 (6)	3.5 (4)	3.2 (6)	1.8 (1)
<i>Missing</i>	7.3 (28)	6.2 (7)	17.2 (32)	14.5 (8)

Employment status (% (n))

<i>No employment</i>	22.7 (87)	25.7 (29)	32.8 (61)	40.0 (22)
<i>Minor employment</i>	3.1 (12)	4.4 (5)	3.2 (6)	1.8 (1)
<i>Part-time employment</i>	20.9 (80)	16.8 (19)	14.0 (26)	14.5 (8)
<i>Full-time employment</i>	44.4 (170)	38.9 (44)	35.5 (66)	34.5 (19)
<i>Missing</i>	8.9 (34)	14.2 (16)	14.5 (27)	9.1 (5)

Living in partnership (% (n))

<i>Missing</i>	7.0 (27)	8.0 (9)	14.0 (26)	9.1 (5)
----------------	----------	---------	-----------	---------

PHQ-9 at baseline (M (SD))

₁SCM = Stepped collaborative care model (intervention group); ₂TAU = Treatment as usual (control group); ₃German: Hauptschule (9 years of education); ₄German: Realschule (10 years); ₅German: Gymnasium (13 years).

Table 2: Results of the primary and sensitivity analyses on differential treatment outcome of SCM₁ at T3.

Mild-Moderate Depression (PHQ-9 <15)							Severe Depression (PHQ-9 ≥ 15)					Group x se- verity degree inter- action		
SCM		TAU ₂		Be- tween group differ- ence	Be- tween group differ- ence	Effect size (Co- hen's d)	SCM		TAU		Be- tween group differ- ence		Be- tween groups differ- ence: p- value	Effect size (Co- hen's d)
Ad- just- ed mean chang e	95 % CI	Ad- just- ed mean chang e	95 % CI	Ad- justed mean (95% CI)			Ad- just- ed mean chang e	95 % CI	Ad- just- ed mean chang e	95 % CI	Ad- justed mean (95% CI)			

Replacement of missing values with Last-observation-carried-forward

-6.23	-	-2.36	-	-3.87	<0.00	0.62	-2.6	-	-1.04	-	-1.56	0.093	0.25	0.036
7.0			3.5	(-5.13;	1			3.6		2.7	(-3.38;			
3;-			3;-	-2.61)				1;-		0;	0.26)			
5.4			1.1					1.6		0.6				
3			9					0		15				

Analysis of complete cases

-7.89	-	-3.41	-	-4.48	<0.00	0.61	-5	-	-2	-	-3.00	0.028	0.47	0.339
9.0			5.0	(-6.12;	1			6.7		4.3	(-5.67;			
8;-			6;-	-2.84)				2;-		4;	-0.33)			
6.7			1.7					3.2		0.3				
0			6					9		3				

Replacement of missing values with Multiple Imputation

-7.07	-	-3.01	-	-4.05	<0.00	0.66	-4.43	-	-1.48	-	-2.95	0.004	0.47	0.332
7.7			4.1	(-5.39;	1			5.5		3.1	(-4.92;			
5;-			8;-	-2.72)				0;-		6;	-0.97)			
6.3			1.8					3.3		0.1				
8			5					6		9				

Replacement of missing values with pattern mixture model

-7.42	-	-2.92	-	-4.50	<0.00	0.51	-5.3	-	-1.76	-	-3.54	0.001	0.49	0.34
8.6			4.4	(-5.71;	1			6.8		3.8	(-5.60;			
9;-			9;-	-3.30)				3;-		3;	-1.48)			
6.1			1.3					3.7		0.3				
6			4					6		1				

¹SCM: Stepped and collaborative care; ²TAU: Treatment as usual.

Table 3: Treatment utilization for different patient groups in SCM₁ or TAU₂ within 12-month study participation

Intervention % (n) [confidence interval]	Total sample		MMD ₅ (PHQ-9 < 15)		SD ₆ (PHQ-9 ≥ 15)	
	SCM n=349	TAU n=98	SCM n=270	TAU n=79	SCM n=69	TAU n=29
No intervention	1.5 (5) [0;3]	0.9 (1) [0;5]	1.9 (5) [1;4]	1.3 (1) [0;7]	0.0 (0) [0;5]	0.0 (0) [0;12]
GP only	3.5 (12) [2;6]	20.4 (22) [13;29]	3.7 (10) [2;7]	25.3 (20) [16;36]	2.9 (2) [0;10]	6.9 (2) [1;23]
GP consultation (alone or in com-	90.6 (307)	96.3 (104)	90.4 (244)	96.2 (76)	91.3 (63)	96.6 (28)

combination)	[87;93]	[91;99]	[86;94]	[89;99]	[82;97]	[82;100]
Psychotherapy (alone or in combination)	63.4 (215) [58;69]	50.0 (54) [40;60]	62.2 (168) [56;68]	45.6 (36) [34;57]	68.1 (47) [56;79]	62.1 (18) [42;79]
Treatment with antidepressant pharmacotherapy (alone or in combination)	54.6 (185) [49;60]	62.0 (67) [52;71]	49.6 (134) [44;56]	54.4 (43) [43;66]	73.9 (51) [62;84]	82.8 (24) [64;94]
Antidepressant pharmacotherapy only	6.2 (21) [4;9]	19.4 (21) [12;28]	5.9 (16) [3;9]	20.3 (16) [12;31]	7.2 (5) [2;16]	17.2 (5) [6;36]
Psychotherapy only	11.8 (40) [9;16]	9.3 (10) [5;16]	13.0 (35) [9;18]	10.1 (8) [4;19]	7.2 (5) [2;16]	6.9 (2) [1;23]
Psychotherapy and antidepressant pharmacotherapy	39.2 (133) [34;45]	36.1 (39) [27;46]	35.9 (97) [30;42]	29.1 (23) [19;40]	52.2 (36) [40;64]	55.2 (16) [36;74]
Low-intensity treatment (telephone psychotherapy, internet self-help, bibliotherapy) (alone or in combination)	41.0 (139) [36;46]	n.a. ₃	43.7 (118) [38;50]	n.a. ₃	30.4 (21) [20;43]	n.a. ₃
Low-intensity treatment only	19.8 (67) [16;24]	n.a. ₃	20.7 (56) [16;26]	n.a. ₃	15.9 (11) [8;27]	n.a. ₃
Any outpatient depression-specific treatment, ₄	94.7 (321) [92;97]	77.8 (84) [69;85]	94.4 (255) [91;97]	72.2 (57) [61;82]	95.7 (66) [88;99]	93.1 (27) [77;99]
Inpatient in total	15.0 (51) [11;19]	25.0 (27) [17;34]	8.9 (24) [6;13]	21.5 (17) [13;32]	39.1 (27) [28;52]	34.5 (10) [18;54]
No information (missing data)	23 (128)	20 (33)	12 (46)	14 (16)	44 (82)	31 (17)

₁SCM: Stepped and collaborative care; ₂TAU: Treatment as usual; ₃n.a.: not assessed; ₄Any outpatient depression-specific treatment: low-intensity and/or psychotherapy and/or antidepressant pharmacotherapy; ₅MMD: mildly to moderately depressed patients; ₆SD: severely depressed patients; **bold numbers**: confidence intervals that do not overlap between SCM and TAU within the total sample and within each severity group.